

REF: 1
 APT: 1
 18-JUN-1999
 18-JUN-1999

18-JUN-1999

REFERENCE: 1. (bases 1 to 1617)
 AUTHORS: Pradiyaya, A.P., Lee, J.H. and TeJong, J.
 TITLE: Identification of a general transcription factor TFIIAalpha/beta homolog selectively expressed in testis
 J. Biol. Chem. 274 (25), 18040-18048 (1999)
 MEDLINE: 99292779
 REFERENCE: 2. (bases 1 to 1617)
 AUTHORS: Pradiyaya, A.P. and TeJong, J.
 TITLE: Direct Submission
 JOURNAL: Submitted 11-DEC-1999 Molecular and Cell Biology, The University of Texas at Dallas, 2601 N. Floyd Rd., Richardson, TX 75080, USA
 FEATURES: Location, qualifiers
 name: 1..1617
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="testis"
 gene: 1..1617
 /gene="ALF"
 CDS: 16..1452
 /gene="ALF"
 /codon_start=1
 /product="TFIIA large subunit isoform ALF"
 /protein_id="AAD39634.1"
 /db_xref="GI:5091688"
 /translation="MACLNVPVKLYRSVIEDVIEGVNLF AE EGIEEQVLKDLKQLWE
 TKVLQSKATEDFFRNSIQSPLFTLQLPHSLHQTLSSTASLVI PAGRTLPSFTTAE LG
 TSNSSANFTFFGYPIHVPAGVTLQTVSGHLYKVNVPIMVTETSGRAGILQHPIQQVEFQ
 QLGQPSVIQTSVPQLNPWSLQATTEKSQRIETVLQQPAILPSGPVDRKHLENATSDIL
 WSPGNEHKIVFEALLCHQESSHYISLPGVVFPQVSQTNSDVESVLSGSASMAQNLHD
 ESLSTSPHGALHQHNTDQLHLKKNRMYGCDSVKQPRNIHEPSNIPVSEKDCNSQVDE
 SIRVTDDLIGELIQVDGSGDTSSNEEIGSTRDADENEFLGNTDGGDILKVFEEEAADSLC
 NEDSATNSSPNSEPHYINIVEEFLNSGDTVSEIHWPELEFETDNVIVGQYDKIHRFSKNK
 WKPYLKI GVMFEGGFQYVFAKAIGFAEW"
 BASE COUNT: 103 A 30 C 243 G 45 T
 ORIGIN:

Query Match: 100.0 ; Score 1617; DB 39; Length 1617;
 Best Local Similarity: 100.0 ; Prod. No. 1;
 Matches 1617; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

20
 40
 60
 80
 100
 120
 140
 160

[illegible]

1b	1261	TTAGATCAAGGATACTATCTGATGATGATGATTTTAAATTAAGATTCTAGAAAGGATTTT	1270
Qy	1261	ttgaatttgttgagatgatgtttagtgaacaggatgtgcaagacctgtttgacacggataaat	1320
1b	1281	TTAAATTTCTAGATGATGTTATGTGAATAGGATCTGCTGAAACCTGTTTTGACACGGATAAT	1320
Qy	1281	gtaatttgttgtcaatgatgataagatttatcgaaagcaagaacaaatggaaatttgtatttg	1380
1b	1301	GTAAATTTCTGTCACTATGATGATAAGTTTCTGGAATGAAGAACTAATGAAATTTTATTTC	1380
Qy	1301	aaagatgtgtgttatctcttttttggatgagagatcttatatttgcaaaagccattgggtgtt	1440
1b	1321	AAGGATGCTCTTATGCTGTTTTGAGGAGAGACTATGTATTTTGCAAAAGCCATTGGGTGAT	1440
Qy	1441	gagagatggtgaaccttcttgaggtcagtagatctatcttgtgaacatcagttggactata	1500
1b	1441	GCAGAGTGGTAAACCTTGTTGAGGCTCACTACATCTATTTTGTGAACATCAGTTGGACTATA	1500
Qy	1501	ttgcataatgtggaattcattttttatttttgaatatagtcacagcacagagctgtttcaaatttt	1560
1b	1501	TTGCATATTGTGAATTCATTTTTATTTTGAATATAGTCCAGCACAGAGCTGTTCAAATTTT	1560
Qy	1561	ttagttccactgtatggaatttaataaaaattataaattcagatgcagatacaatttacac	1617
1b	1561	TTAGTTCACTGTATGGAATTTAATAAAAATTATAAATTCAGATGCAGATACAATTACAC	1617

OK NINE "3:30-6:00"

translation- "MSTNPKWTFTEHPFAQSSQKSKFHLNQGVCRPNGLKLN
 LKREFFSCSSSTSTFLSLIIVFYFEGHISNDELSTFKDPIGFYCPKACTHVL
 YLIPASSNCHLAISSGESSLITRETELAHALLSUNECTHETIKVGLPLHAINQQA
 ENLHJLQFQFVFAFAWAGHKKLEKIDHETLIPSSKKEFSKNEKMLI
 KSLIKOSLMIYICKKFLHQSAENQSLSLSMHMCANENASSFVPHETLFRSQKSGW
 SFMLKPEKNNMNERWGLILFLKVLGOTLIMYEQGLEKPFKEILQLEFYLSEK
 VNFVSAGKIHTVKLEHVSYTEKRKYHSKIEMVEHPDILQMLKLGSTSYHDFILDELT
 VREFLMKLPAVSKIERNYEEJETSLEINDNFWGNTKCEKPFVLSAVITQIYOLCFVNG
 NLEOFITLNLLELHFRDESYYEKDSEKKGIDILEYHFEHCNVVQEFEQSRIIEFVPLD
 AKRFEMLMRKFTLYNDNLPSFLKNNVWQCAVLEQAFNMAELAQRSSYAGSLRSCD
 NTRIHPVPVPSQWIKALWMTMNLQRQKSLKAFMNRFACLGSLQELESEHPVIQVTVGSAKY
 ESAYLAVVWKIDRLFDKNSSLHFLHOLSYLHLGLSDQELPSDWYFFATVQFSFDTCA
 STETURLGVESDVQFQKHVQQLQACYNIQPKLYFVIEDNEVRNLFABEGLHQVIL
 KDLKQWTFKVLQVATERENFEKNDLFTLQIPASLHQTDSSTASLILPAGTLE
 SFTTAEGLGTSNSSANFTEFGYPILHPAGVILQTVSGHLYKVNPLIMVTETSGFAGILQ
 HFIQVVEQQLGQFSLVLTSTSVFQINPWSLQVTTFSQRTETVLQPAIILFGVDRKHL
 ENATSDILVSPGNEEKIVPEALICHQESSHYISIPGVVFSQVWSQTSNDSVESVLGSA
 SMAQNLHDESLSSTPHGALHQHVTDIQLHLRNEMYGCDSEVKQPRNIEEPSNIIVSEK
 DSNQVQVLSIRVTDIDIGEIIQVDSGSDTSNEEIGSTFEDADENEFLGNIIDGCDLKVP
 EEFASQISNEDSATKSSDNEDPQVNIVEEPLNSGDGVSEQDNPOLFOTDNVIVQYD
 KILHSKNKWKKEYLEKDGVMCEGGRDYVFAAIGDAEW"

BASE COUNT	1149 a	302 c	811 d	1062 e
------------	--------	-------	-------	--------

ORIGIN

Query Match 97.4%; Score 1567.2; FR 39; Length 3624;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1587; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

[illegible]

Qy 79 gntggjaat taa ttgttgaagaggp ataga gaa taa tttaaagantt gaagcag 138
 GTCGGAATCTATTTCGTGAAGAAGCTATAGAGGAAATACTTTTAAAGACTTGAAGCAG
 Sp 1290 2049

QY 159 CTTGGGGAACCAAGGTCTTGCGAGTGTAAAGCAACA GAAACTTCTCTTCAGAAATAGCATC 198
 ||| | | | | | | | | | | | | | | | | | | | |
 DP 235A CTTGGGGAACCAAGGTTTTGCATTCTTAAGGCAACH GA AACTTCTTCAGAAATAGCATC 240

[illegible]

Qy	1394	tttgaagaggaagagactatgatatggcaaaaacacattcttgatgagagtggttaaaacatttg	1458
Nb	4610	TTTTGAGGGAAGAGATTATGATTTTAAAAATATTTCTTAACTAGAGACTATTAAGATTTT	5174
Qy	1414	tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt	1504
Nb	5174	TTTATTTTATATATGATTTTATGAGATTAATTTATATATATATATATATATATGATTTTCA	5714
Qy	1519	tttttatttttgaatatatgctccagacagagctgttcaaaattttttagttccactgtatggaa	1578
Nb	5714	TTTTTATTATATATATAAGCAATGAGAGCTATTGATTTTACTTCAATGATATGAA	6144
Qy	1579	tttaataaaaattataattcagatcagatataatt	1613
Nb	6144	TTTTAAAGAAATATATTTAGATAGAGTATATTT	6284

RESULT 3

AC023827/c

LOCUS

AC023827 151767 bp DNA HTG 09-MAR-2000

DEFINITION

Homo sapiens chromosome 16 clone RP11-42118, WORKING DRAFT
SEQUENCE, 3 unordered pieces.

ACCESSION

AC023827

VERSION

AC023827.2 GI:2211923

KEYWORDS

HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 151767)

AUTHORS

DOE Joint Genome Institute.

TITLE

Sequencing of Human Chromosome 16

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 151767)

AUTHORS

DOE Joint Genome Institute.

TITLE

Direct Submission

JOURNAL

Submitted (18-FEB-2000, Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT

On Mar 9, 2000 this sequence version replaced gi:6997147.

-----Genome Center

Center: Joint Genome Institute

Center Code: JGI

Web site: <http://www.jgi.doe.gov>

-----Library Statistics

Library quality: 99.99% (1000 bases)

Library quality: 99.99% (1000 bases)

Library quality: 99.99% (1000 bases)

Estimated insert size: 151767; sum-of-squares estimation

Estimated insert size: 151767; agarose-gel estimation

Quality coverage: 4.3x in 100 bases; agarose-gel estimation

Quality coverage: 4.3x in 100 bases; sum-of-squares estimation

Notes: This is a working draft sequence. It currently

contains gaps and may be subject to change.

For more information, please contact the JGI staff.

RECEIVED
HONOLULU, HI

1999 FEB 04 10 10 AM
U.S. DEPARTMENT OF JUSTICE
FEDERAL BUREAU OF INVESTIGATION
WASHINGTON, D.C. 20535

04-FEB-1999


```

RESULT      E
HUMTFIIA-42
LOCUS       HUMTFIIA-42 1498 bp     mRNA             FR1              04-FEB-1999
DEFINITION  Human mRNA for TFIIA-42, complete cds.
ACCESSION   D14887
VERSION     D14887.1 GI:727196
KEYWORDS    TFIIA-42; transcription factor.
SOURCE      Homo sapiens cell_line:HeLa-S3 cDNA to mRNA, clone_lib:lambd
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1549)
AUTHORS     Handa,H.
TITLE       Direct Submission.
JOURNAL     Submitted (05-APR-1993) to the DDBJ/EMBL/GenBank databases. Hiroshi
            Handa, Tokyo Institute of Technology, Faculty of Bioscience and
            Biotechnology, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa
            226-8501, Japan (Tel:045-922-1111, Fax:045-923-0380)
REFERENCE   2 (bases 355 to 1549)
AUTHORS     Watanabe,H., Ogari,K., Wada,T., Imai,T., Shiroya,T. and Handa,H.
TITLE       Isolation of two cDNAs that encode the large subunit of human
            transcription factor IIA
JOURNAL     Unpublished (1993)
FEATURES    Location/Qualifiers
            source          1..1549
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /cell_line="HeLa-S3"
                        /clone_lib="lambda gt10"
CDS         401..1535
            /note="(seq:435..1549) is same as in D14886; major
            /translational_start="
            /cdon_start=1
            /product="TFIIA-42"
            /protein_id="BAA03604.1"
            /db_xref="GI:727197"
            /translation="MANSANTNTVPELYRSVIEDVINVDVFIDLDGQVDEQVLMELKT
            LWENKLMCSRAVDGFHSEEQQLLLQVQQHQPPQQQHHHHHHHHQQAQPPQTVPQQAQT
            QQVLIPASQQATAPQVIVFD SKLIQH MNASNMSAAATAATLALPAGVTPVQQILTNSG
            QLLQVVRAANGAOYIFQPQQSVVLQQQVI PQMQPGGVQAQFVIQQVLAPLPSSISPTGT
            VILQPPQQLFTGNKIQVIPITVAAPTPAQAQLTATGQQQFAQAQPAQTQAPLVLOVDST
            GETSSEEDDEDEEDYDDDEEDDKKDGAE DGQVEEPLNSEDDVSDEEGQE LFD TENV
            ITNTPYKIKRKKKKPKHIHNNNNNNRIYIIRKAIQDAEK"

```


Qy 1168 gaagaagaaggtgacagtatttcacatgaggatcaggcaaaaacagtatgataatgaa 1177
 Th 847 GAAAGATGAAGATGAAGAAAGAAACTATGATGATGATGAGGAGGAAGACAAAGAGAAAGAT 906
 Qy 1187 gacattgagttacatcttcttgaagaggaacatttcaattcttgaagttaattgttagtga 1197
 Dr 867 GAAAGTTGAAGATGAGCAGTTGAAAGAAAGAGGCTTCAATATGAAGATGATGTGAGTGTAT 906
 Qy 1255 gaggatggtgagagctgtttgagatgataaattgtaattgtctgtcagtatgataaqatt 1347
 Db 967 GAGGAAGGACAGCAACTCTTTGACACAGAAATGTGTGTGTATGCAATATGATAAGATA 1026
 Qy 1345 catcgaaggaagaacaaatggaaattcaattgaagatgggtgttatgtgttttggaggg 1407
 Db 1027 CACAGAAGTAAAAACAAATGGAAATTTTCATCTCAAGCAAGGCATTATGAATCTTAATGGA 1086
 Qy 1405 agagactatctatcttgcacaaagcatttggatgagagagtgttaa 1452
 Db 1087 AGAGATTATATATTTTCCAAAGCCATTGCAGATGCACAATGCTCA 1131